Scientia Iranica E (2011) 18 (3), 722–730

Sharif University of Technology

Scientia Iranica

Transactions E: Industrial Engineering

www.sciencedirect.com

An enhanced neural network model for predictive control of granule quality characteristics

N. Neshat [a], **H. Mahlooji** [b,*], **A. Kazemi** [c]

[a] *Department of Industrial Engineering, Faculty of Engineering, Tarbiat Modares University, Tehran, P.O. Box 14155-143, Iran*

[b] *Department of Industrial Engineering, Sharif University of Technology, Tehran, P.O. Box 11155-9414, Iran*

[c] *Department of Industrial Management, Faculty of Management, University of Tehran, Tehran, P.O. Box 14155-6311, Iran*

Received 3 March 2010; revised 29 September 2010; accepted 6 December 2010

KEYWORDS

Predictive control;
Artificial neural networks;
Partial correlation;
Particle size;
Modelling;
Spray drying.

Abstract An integrated approach is presented for predicting granule particle size using Partial Correlation (PC) analysis and Artificial Neural Networks (ANNs). In this approach, the proposed model is an abstract form from the ANN model, which intends to reduce model complexity via reducing the dimension of the input set and consequently improving the generalization capability of the model. This study involves comparing the capability of the proposed model in predicting granule particle size with those obtained from ANN and Multi Linear Regression models, with respect to some indicators. The numerical results confirm the superiority of the proposed model over the others in the prediction of granule particle size. In order to develop a predictive-control strategy, by employing the proposed model, several scenarios are developed to identify the most suitable process settings with respect to the desired process response. Utilization of these scenarios paves the way for decisions about spray drying to be made consistently and correctly without any need for judgmental speculations or expensive trial-and-error tests.

1. Introduction

For over a century, spray drying, as a unique drying process for powder production, has experienced extended applications in the foodstuff, manufacturing and pharmaceutical industries.

A simplified description of the granule production process is as follows (see Figure 1): The slip with specified density and viscosity is fed at a constant pressure and temperature by the pump (1), passes through the filters (2), reaches the distributor ring (3), inside the tower (4). The finely atomized slip jet is hit by a vortex of hot air at a fixed temperature produced by natural gas (8), and an intake fan (7). The air, which is conveyed to the upper part of the tower via a heat-insulated steel duct, causes a constant suction pressure in the tower (9) and is set in rotation by the annular distributor (10). Dried powder, i.e. the granule, is unloaded via the outlet valve (5) onto a conveyor belt. The fine powder residue that remains suspended in the air is exhausted by the main fan (11) and is separated out, in part, by the cyclones (6) and, in part, by the wet dust separator (12). Then the exhausted air, at a constant temperature, leaves via the stack (13) [1].

Granule particle (PS) size, as a quality characteristic, is controlled by measurement during granule production. The Weight of the Residual Granule (WRG) is the weight of the output granule with a diameter greater than 300 μm in a sample that weighs 0.1 kg. Hence diameters must be monitored in order to control the PS quality. According to the spray drying process diagram shown in Figure 2, the process response, i.e. WRG, is considered in terms of its related variables. Such variables usually have to do with either properties or operating conditions, like slip viscosity (ν) and density (ρ). There are also a few in-process variables, such as the hot air temperature (T_{in}), air suction pressure (P_S) and fed slip pressure (P_p), to be considered.

The exhausted air temperature (T_{out}), as an in-process variable, is generated through internal treatment of the spray drying process [1]. A few features of the input, in-process and output variables of spray drying are summarized in Table 1.

* Corresponding author.
E-mail address: mahlooji@sharif.edu (H. Mahlooji).

doi:10.1016/j.scient.2011.05.019

Production and hosting by Elsevier

Nomenclature

ANFIS	Adaptive Neuro-Fuzzy Inference System
PSI	Particle Size Index
ANN	Artificial Neural Network
R^2	absolute fraction of variance
MAPE	Mean Absolute Percentage Error
RMSE	Root Mean Square Error
MFs	Membership Function
T_{in}	inlet air temperature
NN	Neural Network
T_{out}	exhausted air temperature
PLS	Partial Least Squares
WRG	Weight of Residual Granule
P_p	fed slip pressure
y_{dm}	target value of output
P_S	air suction pressure
y_m	obtained value of output

Greek symbols

ν	viscosity
ρ	density
γ	learning rate
$\Delta\psi_m$	bias values of output layer to hidden layer
$\Delta\psi_j$	bias values of input layer to hidden layers
ΔW_{ij}	weighted value revising function between input and hidden layer
ΔW_{jm}	weighted value revising function between hidden and output layers

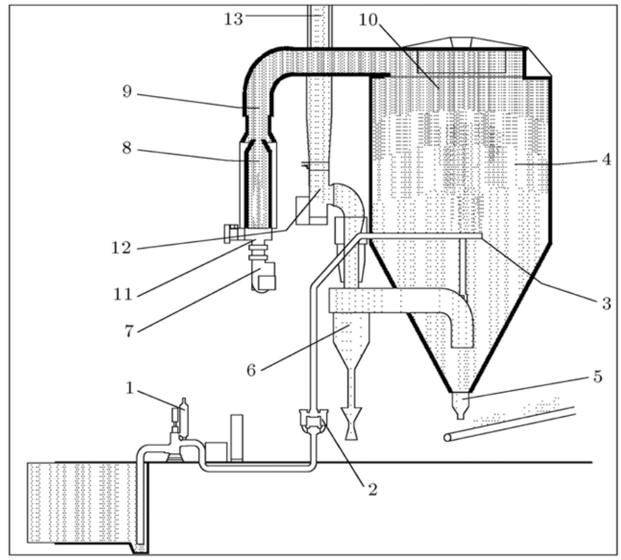

Figure 1: Schematic diagram of a spray drier.

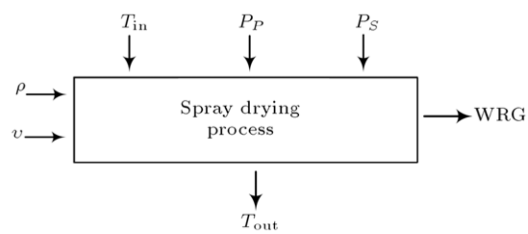

Figure 2: Process diagram of spray drying.

Often, a full theoretical understanding of the mechanism of a complex process such as spray drying may be lacking [2]. The lack of such understanding is directly attributed to the complexity of the interdependent and correlated process variables. Constructing a model of the process is a conventional alternative for discovering the relationships between input and output variables of the process with the aim of controlling the process.

To judge by recent publications, conventional parametric methods are not suitable for modelling complex manufacturing processes [3–5] due to the great number of process variables and the non-linear nature of their relations. In contrast, Artificial Neural Networks (ANN) have been known to be good candidate approaches for this purpose. Several successful implementations of ANN have been reported recently. Su and Hsieh [6] compared the reliability of the Taguchi model with the ANN model for semiconductor manufacturing processes and established the superiority of the ANN model. Yimnirun et al. [7] studied the modelling of the synthesis process of $N_iNb_2O_6$ powder using ANN and multiple regression models. In their investigation, ANN significantly outperformed Multiple Regression. Furthermore, several studies investigated the process mechanism and optimization of the manufacturing process through neural network modelling [8–13].

Neural network modelling is an area with considerable potential for improving quality control processes in the ceramics industry as well. The literature on such applications is more recent and less extensive. For instance, Lam et al. [14] developed a prediction module to estimate two quality metrics of slip-cast pieces through simultaneous execution of two neural networks and a process improvement algorithm. Their algorithm optimizes controllable process settings using the neural network prediction module in the objective function. Scott et al. [15] developed an ANN model in order to use compositional information to predict the relative permittivity and oxygen diffusion properties of ceramic materials. The relevant results show that ANN is able to produce accurate predictions of the properties of these ceramic materials.

Table 1: Characteristics of proposed variables of spray drying process.

	Units	Type	Interval		Mean value	Standard deviation
			Min	Max		
Slip density	kg^{-3}	Input	1.66×10^3	1.70×10^3	1.68×10^3	0.01×10^3
Slip viscosity	S	Input	45.00	70.00	57.02	9.04
Inlet air temperature	°C	In-process	540.00	555.00	549.12	2.94
Fed slip pressure	N/m^2	In-process	13.50×10^5	18.50×10^5	17.30×10^5	0.70×10^5
Air suction pressure	N/m^2	In-process	18.50×10^2	21.00×10^2	18.54×10^2	1.30×10^2
Exhausted air temperature	°C	In-process	91.00	104.00	98.27	2.64
Weight of residual granule	kg	Output	68.10×10^{-3}	77.00×10^{-3}	74.22×10^{-3}	3.68×10^{-3}

Dinh and Afzulpurkar [16] compared ANN and the Co-Active Neuro-Fuzzy Inference System (CANFIS) in modeling the real complicated Multi-Input–Multi-Output (MIMO) nonlinear temperature process of a roller kiln used in the ceramic tile manufacturing line. A three-layer Back Propagation Artificial Neural Network (BPANN) was developed by Yu et al. [17] for predicting the performance of porous Si_3N_4 ceramics. The results indicated that BPANN is a very useful and accurate tool for the prediction and optimization of porous Si_3N_4 ceramics performance. However, earlier studies can be found for this process in other industries [18–20] also. Since the most popular application of ANN to spray drying is the prediction of process parameters, citing the following works is in order: Chegini et al. [21] predicted seven performance indices in an orange juice spray drier using ANN. ANN and Response Surface Methodology (RSM) were compared by Youssefi et al. [2] in predicting the quality parameters of spray dried pomegranate juice. The results showed that the ANN model outperforms RSM for complex spray drying.

Ant studies published in the area of modelling complex manufacturing processes can be categorized into two groups. The first group includes those attempts to establish the superiority of the ANN algorithm over parametric methods, such as RSM and Multiple Linear Regression (MLR) analysis. The second group consists of studies using an ANN algorithm or a hybrid system of ANN along with other heuristic algorithms, such as the Genetic Algorithm (GA), with the aim of predictive control and optimization.

In this study, an integrated intelligent approach consisting of Partial Correlation (PC) analysis (parametric approach) and ANN (heuristic approach), code named PC-ANN, is used to develop a reliable and accurate model for predicting the granule PS. According to the proposed algorithm, initially, the PC analysis is applied to process the actual data and provide the necessary background to apply ANN. Based on the findings of the PC analysis, at the next step, the PC-ANN model is developed to predict the granule Particle Size Index (PSI). As a comparative study, the performances of the PC-ANN and ANN models are studied. Finally, based on the superior approach, several scenarios are proposed in order to achieve predictive control of the granule PS as accurate, fast running and inexpensive tools.

The remainder of this paper is organized as follows. Section 2 briefly describes artificial neural networks and modelling with ANN. The proposed PC-ANN algorithm is introduced in Section 3. Section 4 is devoted to experiments and the actual case study of this work. Discussion and scenarios are included in Section 5 and conclusions are presented in Section 6.

2. Modeling with ANN

ANNs mimic the ability of biological neural systems in a computerized way by resorting to the learning mechanism as the basis of human behaviour [22]. Utilizing samples from process outputs, ANN proposes an approximate model architecture to fit the data. Therefore, ANN may be applied to problems with a non-linear nature or with too complex algorithmic solutions. Their ability to perform complex decision-making tasks without prior programming makes ANNs more attractive and powerful than parametric approaches, especially for complex problems.

Whereas the accurate prediction of a quality characteristic is a key factor in the success of a manufacturing operation, ANN can be deployed to model and predict the response of a complex manufacturing process. In order to model the process, the network must be trained based on input and in-process variables, e.g. process settings, as well as process responses/outputs, e.g. quality characteristics. The ANN then 'learns' the governing relationships in the input and output data sets by modifying the weights between its nodes. In essence, a trained ANN model can be considered a function that maps input vectors to output vectors. Briefly, the intention is that the network should behave exactly like the process, considering its response to variables and conditions, hence providing a model of the process which can be used for the prediction of output properties.

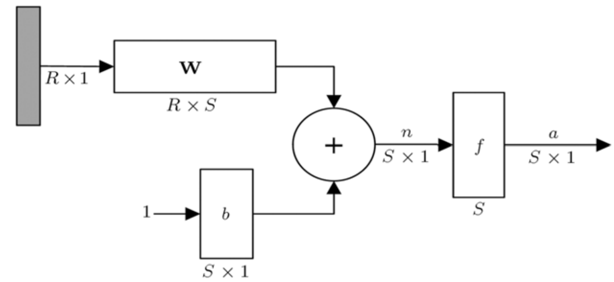

Figure 3: The network architecture of an ANN model consisting of one hidden layer with S neurons, along with input and output layers, with R and one neuron(s).

The architecture of an ANN model usually consists of three parts: an input layer, the hidden layers and an output layer. The information contained in the input layer is mapped to the output layer through the hidden layers. Each neuron can receive its input only from the lower layer, and sends its output to the neurons only on the higher layer.

Figure 3 illustrates the network architecture of an ANN model consisting of one hidden layer with S neurons, along with input and output layers with R and 1 neuron(s), respectively. Here, input vector $\mathbf{p}$ with R input elements is represented by a solid dark vertical bar on the left. Thus these inputs post multiply the S-row, R-column matrix $\mathbf{W}$. The quantity 1, as a constant value, enters the neuron as an input and is multiplied by vector bias b with S rows and a single column. The net input to the transfer function, f, consists of n, the sum of bias b, and the product $\mathbf{W_p}$. This summation is passed to the transfer function, f, to get the neuron's output which in this case is a vector with length S [23].

3. Methodology

Modelling complex and nonlinear manufacturing processes that deal with noisy, limited and non-integrated data requires methods that can alleviate these problems. Thus we propose an integrated intelligent approach, code named PC-ANN, to achieve this purpose. In this study, both ANN and PC-ANN approaches are applied to predict granule PSI, in an attempt to identify the superior approach.

ANN and PC-ANN approaches are applied in order to develop ANN and PC-ANN models for optimum prediction of the granule, PSI. Although both these approaches take a similar set of six phases (as follows) to develop the models, it should be noted that their models are different, with regard to their input set.

These six phases are described as follows:

Phase 1. *Selection of neural network topology*. At present, there are about 31 different Neural Network (NN) topologies being employed in research. The most common networks are: Multi Layer Perceptrons (MLP), also called Multilayer Feed Forward Networks, Adaptive Resonance Theory Models (ART), Recurrent Associative Networks (RAN) and Self-Organizing Maps (SOM) [24]. Each combination of an ANN type and a training algorithm is selected for a different situation, depending on the networks purpose of usage. For instance, the MLP trained with a Back Propagation (BP) algorithm is suitable as a black-box model of systems whose underlying relations are vaguely known or are extremely complex. The most popular algorithm in engineering applications is the standard Back Propagation (BP) algorithm [25]. This algorithm is a widely used iterative optimization technique that locates the minimum of a function expressed as:

$$E = \frac{1}{2}\sum_m (y_{dm} - y_m)^2, \tag{1}$$

where y_{dm} is the target output value of the output layer, and y_m is the obtained value of the output layer. Based on the BP algorithm, during the training process, the deviation between the network output and the desired output at each presentation is computed as an error. This error, in quadratic form, is then fed back (back propagated) to the network and is used for modifying the weights by a gradient descent method, as follows:

$$\Delta\omega_{jm} = -\gamma \frac{\partial E}{\partial W_{jm}} = \gamma \frac{\partial E}{\partial y_m}.\frac{\partial y_m}{\partial \mathrm{net}_m}.\frac{\partial \mathrm{net}_m}{\partial W_{jm}}, \tag{2}$$

$$\frac{\partial E}{\partial y_m} = \frac{\partial}{\partial y_m}\left(\frac{1}{2}\sum_m (y_{dm} - y_m)^2\right) = -(y_{dm} - y_m), \tag{3}$$

$$\frac{\partial y_m}{\partial \mathrm{net}_m} = f_o'(\mathrm{net}_m), \tag{4}$$

$$\frac{\partial \mathrm{net}_m}{\partial W_{jm}} = \frac{\partial}{\partial W_{jm}}\left(\sum_m P_j W_{jm}\right) = P_j, \tag{5}$$

$$\Delta W_{jm} = -\gamma[-(y_{dm} - y_m)]f_o'(\mathrm{net}_m)P_j = -\gamma\theta_m P_j, \tag{6}$$

$$\theta_m = -(y_{dm} - y_m)f_0'(\mathrm{net}_m), \tag{7}$$

$$\Delta\psi_m = -\gamma\theta_m. \tag{8}$$

The weight change of the hidden layer to the input layer is calculated s follows.

$$\Delta W_{ij} = -\gamma\theta_j X_i, \tag{9}$$

$$\begin{aligned}\theta_j &= \sum_m \frac{\partial E}{\partial P_j}\frac{\partial P_j}{\partial \mathrm{net}_j} = \sum_m \frac{\partial E}{\partial y_m}\frac{\partial y_m}{\partial \mathrm{net}_j}\frac{\partial \mathrm{net}_j}{\partial P_j}\frac{\partial P_j}{\partial \mathrm{net}_j} \\ &\quad \times \sum_m [-(y_{dm} - y_m)] f_0'(\mathrm{net}_m) \\ &\quad \times \frac{\partial}{\partial P_j}\left(\sum_j P_j W_{jm}\right)\frac{\partial P_j}{\partial \mathrm{net}_j} \\ &= -\sum_m (y_{dm} - y_m) f_0(\mathrm{net}_m) W_{jm} f_P'(\mathrm{net}_j) \\ &= \sum_m (\theta_m W_{jm}) f_P'(\mathrm{net}_j), \end{aligned} \tag{10}$$

$$\Delta W_{ij} = -\gamma \sum_m (\theta_m W_{jm}) f_P'(\mathrm{net}_j) X_i = -\gamma\theta_j X_i, \tag{11}$$

$$\Delta\psi_j = -\gamma\theta_j, \tag{12}$$

where γ is the learning rate, $\Delta\psi_m$ and $\Delta\psi_j$ stand for the bias values of the output layer to the hidden layer, as well as the input layer to the hidden layers, respectively, and ΔW_{ij} and ΔW_{jm} are the weighted value revising function between the input layer and the hidden layer, as well as between the hidden layer and the output layer, respectively.

In this phase, the topologies of the network and its training algorithm, considering the nature of modelling the problem, are chosen. The other decision to be made is defining how the input variables and the process response are presented to the network as input and output, respectively. One way is to present all available process variables as network input, and then let the network modify itself during training so that the connection of any insignificant variables becomes weak. Another approach that is more selective introduces as input only those variables that are surely affecting the process output. The first approach has been called the "global approach" while the second is termed the "focused approach" [26].

Phase 2. *Collection and preparation of training data set*. In the data collection stage, it is necessary to ensure the sufficiency and integrity of the data used to train, validate and test the network. As the performance of a network can be directly influenced by the presented data, it is not simply possible to say how many data sets are required, because this depends on the nature of the modelling problem and the cost of providing data. The prepared set of data is partitioned randomly into three sets: training, validation and test data.

Training data can be provided in three ways [27].

1. Simulated data: The data in this category is replaced by actual process data. Often, simulated data is generated by statistical models (such as normal distribution) and computational simulations (such as finite-element analysis methods).
2. Actual process data: When randomly selected raw process data is used for training NN, in fact, the actual data is employed to construct NN model. Many manufacturing companies have already provided both process variable data and quality inspection/product test data in their databases.
3. Designed experiment data: This data is not collected directly from normal production conditions, but is obtained from designed experiments on the process, often using the Taguchi or Design Of Experiment (DOE) approach.

In order to enhance model fitness, several tasks for preparing data may be performed, such as:

1. Data integrity check,
2. Extreme data removal,
3. Data pre-processing and post-processing,
4. Data coding.

Phase 3. *Constructing and fitting the neural network model*. The performance of the NN model depends on the fitness of the network features. For instance, too few neurons may result in under-fitting, but too many neurons may yield over-fitting, which means that all the training data fit well, but the NN model performance for the test data is mediocre. The optimal configuration of the network is selected according to the value of the training error function. So, various architectures with different features of neural networks can be proposed. During the training process, representative examples of inputs and their corresponding outputs of process are presented to the networks. Ultimately, the best network is chosen based on the user-specified training error function.

Phase 4. *Model validation*. The validation data set is used to ensure that there is no over-fitting in the final result. In order to validate the model, a data set is randomly selected from the training data. When a significant over-fitting occurs, the error of

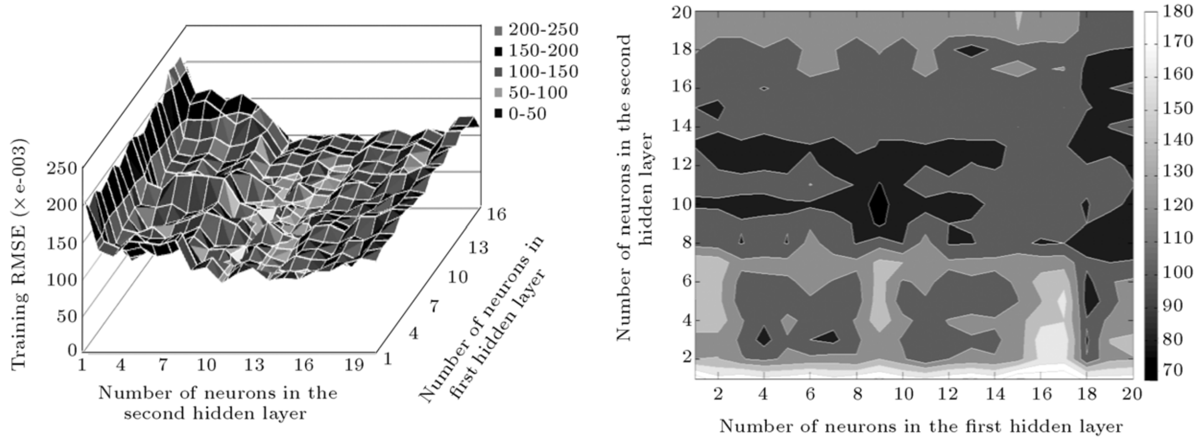

Figure 4: Effects of changing the number of neurons in the first and second hidden layers on the learning performance of the ANN model.

validation data starts to increase and the training process comes to an end.

Phase 5. *Selection of performance indicators*. To determine the reliability of a model, several performance indicators can be used. The performance of the ANN model, based on model reliability, is evaluated by a regression analysis between the predicted and actual values.

The indicators used for evaluation of the network performance include mean relative error, root mean square error and coefficient of determination [28]. The mean relative error which shows the mean ratio between the error and the actual values is calculated as:

$$\text{MRE}\ (\%) = \frac{1}{m}\sum_{i=1}^{m}\left|100\frac{(y_i-\hat{y}_i)}{y_i}\right|, \tag{13}$$

where $\hat{y}_i$ is the estimated value by the ANN model, y_i is the actual value of the response process, and m is the number of points in the data set. The root mean square error is calculated as:

$$\text{RMSE} = \sqrt{\frac{1}{m}\sum_{i=1}^{m}(y_i-\hat{y}_i)^2}. \tag{14}$$

Finally, the coefficient of determination, a statistical criterion that can be applied to multiple regression analysis, is calculated as:

$$R^2 = 1-\left(\frac{\sum_{i=1}^{m}(y_i-\hat{y}_i)^2}{\sum_{i=1}^{m}(y_i-\bar{y})^2}\right). \tag{15}$$

The coefficient of determination ranges between zero and one. Ideally, R^2 should be close to one, whereas a poor fit results in a value near zero.

Phase 6. *Model performance evaluation*. By termination of process training and validation, the network is ready for prediction. Hence, input vectors from separate test data are introduced to the trained network, and the responses of the network, i.e. the predicted outputs, are compared with the actual ones using the performance indicators.

The significance of the integrated PC-ANN approach for optimum prediction of the granule PSI is threefold. First, it normalizes all input and output values by pre-processing and post-processing, respectively, to eliminate all possible noise. Second, it is capable of handling the complexity of relations, limitations and noise in the given data. Third, it improves the predictive capability of the neural network via dimension reduction operation.

4. The experiment

In order to develop the ANN and PC-ANN models for predicting the granule PSI, a set of actual data consisting of 300 data vectors was collected under steady state conditions from a real system. This data was collected from a large ceramic tile factory in Yazd province in south Iran. The collected raw data was provided randomly from the production database for the period March 23rd, 2005 to September 22nd, 2006 (see the Appendix). Then, 80% of the data set was randomly set aside for training purposes, 10% as the validation, and the remaining 10% for testing the network. All the input and output values were normalized by pre-processing, and checked for integrity by histogram plots.

4.1. The ANN model

Taking a global approach in presenting data to the network, the ANN model adopted for predicting the granule PSI was MLP, trained by the BP algorithm, based on the Levenberg–Marquart rule. The network includes five inputs, which consist of ρ, ν, P_P, P_S and T_{in}, while it has one output, which is WRG.

For choosing the optimal configuration of ANN for predicting the granule PS via predicting the WRG, a number of different network configurations consisting of one to three hidden layers and different numbers of neurons in hidden layers with various transfer functions were considered. The training process was run in a MATLAB environment, so the minimum of a user-specified error function, i.e. RMSE, was reached, while the number of epochs (frequency of presenting training data to the network) was less than 200, and the grad rate (deviation of training error in each epoch) was greater than 1e−010. The optimal configuration of each network was selected based on values of its training and validation RMSE. As Figure 4 indicates, when the numbers of neurons in the first and second hidden layers are equal to nine and ten, respectively, we get the least training RMSE value for the training process in predicting the granule PSI. However, in this network, selection of more than three layers leads to a decrease in the network's training performance.

The optimum network architecture of the ANN model for predicting the granule PSI is given in Figure 5. This figure shows

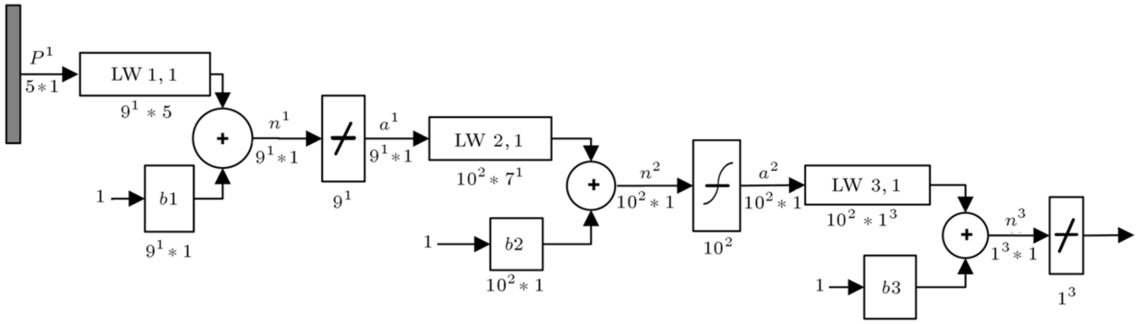

Figure 5: The network architecture of the ANN model for predicting the granule PSI.

Table 2: The detailed information of optimum configuration and the results of performance evaluation of the PC-ANN, ANN and the MLR models for predicting the grannule PSI.

	Optimum configuration	Training		Performance evaluation		
		Training RMSE	Validation RMSE	R^2 (%)	MRE (%)	RMSE (kg)
PC-ANN model	3-6-8-1	5.82358e−005	3.2451e−004	88.16	5.582	1.014×10^{-3}
NN model	6-7-10-1	2.37658e−004	8.10037e−004	82.32	8.945	1.784×10^{-3}
MLR model				68.09	15.67	4.056×10^{-3}

Table 3: The PC analysis between the process response and the other process variables.

	WRG, P_S	WRG, T_{in}	WRG, P_p	WRG, T_{out}	WRG, ν	WRG, ρ
PC coefficient	−0.055	0.111	0.002	0.249	0.567	0.316
P_{value}	0.601	0.497	0.983	0.017	0.000	0.002
Status	Not correlated	Not correlated	Not correlated	Correlated	Correlated	Correlated

that the network including two hidden layers with nine and ten neurons is the preferred network. The activation functions in the layers of this network were chosen as the purline (which generates output the same value of input), tangent sigmoid (which generates output between −1 and 1 as the neuron's net input goes from negative to positive infinity) and purline transfer functions, respectively.

The ANN model for predicting the granule PSI was validated by a stopping rule after 87 epochs when the error due to the validation data started to increase. Detailed information for optimum configuration and the results of the training of the ANN model for predicting the granule PSI are given in Table 2. The ANN model predictions for granule PSI resulted in an MRE of 8.945%, an RMSE of 1.784×10^{-3} kg and an R^2 of 82.32% in comparison to the actual values for the test data set.

4.2. PC analysis

In order to discover how the input and in-process variables influence the process response and discover pair interactions between process variables, a PC analysis was applied to the process variables. The PC produces coefficients that describe the relationship between two variables while adjusting for the effects of one or more additional variable. In fact, the PC coefficient measures the degree of association between two random variables, with the effect of a set of controlling random variables removed [29]. In order to calculate the PC coefficient, a statistical test whose null hypothesis is "two variables are correlated" must be conducted. The null hypothesis will be rejected if the P_{value} (the probability of observing data at least as extreme as that observed) for these tests is smaller than the selected value of the level of significance.

As illustrated in Table 3, using SPSS software, the partial correlation coefficients between the process response and each of the other variables were calculated, and the correlation status of each pair was determined. The results indicate that only ρ, ν and T_{out} are evidently correlated with the process response, considering that the P_{value} for these tests is larger than the selected value of the level of significance (0.025). Evidently, the other four variables, P_P, T_{in}, P_S and T_S, are not significantly correlated to WRG.

Table 4: The PC analysis of pair interactions of input and in-process variables of spray drying.

	P_S, T_{out}	P_P, T_{out}	T_{in}, T_{out}	ν, ρ
PC coefficient	0.31	−0.324	0.354	0.477
P_{value}	0.003	0.035	0.001	0.000

To provide a better insight, the correlation analysis was applied to input and in-process variables and the results are illustrated in Table 4 (for cases with considerable linear correlation). It should be noted that the availability of nonlinear relations between variables that were not linearly correlated was also investigated. Our investigation failed to register any significant non-linear relation either.

The following conclusions can be drawn, based on the results displayed in Table 4.

1. ρ and ν are positively related, which is reasonable because as the density, i.e. the water percentage of slip, decreases, the viscosity of slip also decreases.
2. P_P, T_{in} and P_S are pair wise independent, while they all are correlated to T_{out}.

The concurrent analysis of the results contained in Tables 2 and 3 implies that P_P, T_{in} and P_S have directly influenced T_{out}, which means that they are indirectly related to WRG. Thus T_{out}, as an

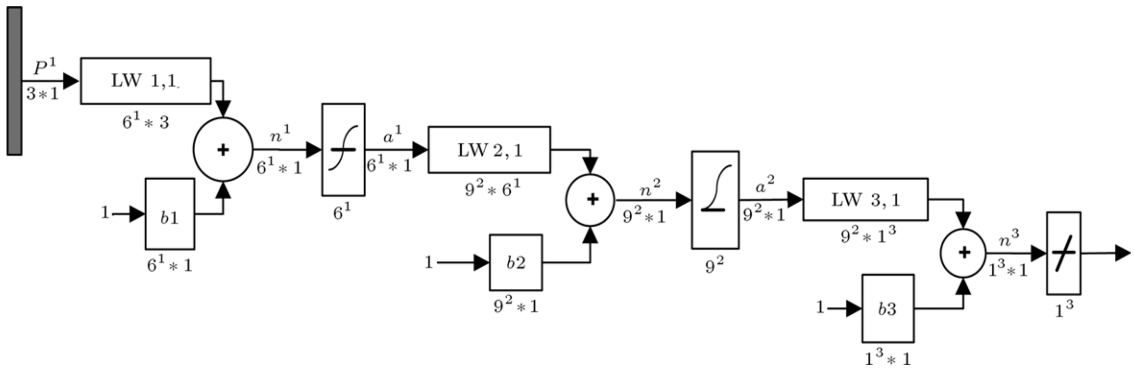

Figure 6: Network architecture of the PC-ANN model for predicting the granule PSI.

explanatory variable for the process response, is dependent on P_P, T_{in} and P_S. Also within the neural network training process, T_{out} can incorporate some additional information in predicting the granule PSI.

4.3. The PC-ANN model

The PC-ANN model was developed considering the findings of the PC analysis presented in the previous section. First, we used ANN to estimate T_{out} in terms of P_P, T_{in} and P_S. Second, we constructed the PC-ANN model for estimating the WRG in terms of the estimated T_{out} as a hint [10], as well as two input variables, i.e. ν and ρ. According to the six stages of ANN modelling, the PC-ANN model was constructed.

In order to compare the performance of the proposed ANN model and the PC-ANN model of spray drying, the same test data vectors were used to evaluate the performance of the PC-ANN model.

To identify the optimal configuration of the PC-ANN model for predicting the granule PSI via predicting WRG, all plausible PC-ANN models were run and estimated with regard to various network features. Based on different numbers of hidden layers and neurons, as well as various transfer functions, a number of different network configurations were considered. The computer code for solving the BP algorithm (MATLAB 7, Release 14, The Mathworks Inc., MA, USA) was implemented. Training information and the results of the PC-ANN model are shown in Table 2.

The training process was run, and led to the minimum value of the RMSE function, while the number of epochs was less than 200 and the grad rate was greater than 1e−010. The optimal configuration of the network was determined based on the value of training and validation, RMSE. The sensitivity analysis of the RMSE value showed that when the number of neurons in the first and second hidden layers are equal to six and nine, respectively, it assumes its smallest value in predicting the granule PSI. However, when the number of hidden layers exceeds three, the performance of the training process starts to deteriorate.

The optimum network architecture of the PC-ANN model for predicting the granule PSI is given in Figure 6. This figure shows that the network consisting of two hidden layers with six and nine neurons is the most preferred network. Activation functions in the layers of this network were chosen as the tan sigmoid, log sigmoid and pure line transfer functions.

The ANN model for predicting the granule PSI was validated by a stopping rule after 240 epochs when the error due to the validation data started to increase. Detailed information about the optimum configuration and the results of training of the PC-ANN model for predicting the granule PSI are given in Table 2. The PC-ANN model predictions for the WRG led to an MRE of 5.582%, an RMSE of 1.014×10^{-3} kg and an R^2 of 88.16%, when compared to the actual values for the set of the test data.

4.4. Multi linear regression model

Multi Linear Regression (MLR) estimates the coefficients of a linear equation involving two or more independent variables that best predict the value of the dependent variable. Considering the fact that T_{out} is correlated to T_{in}, P_S and P_p, the MLR model for predicting the granule PSI via estimating the WRG is illustrated by the following equations:

$$T_{\text{out}} = 0.249T_{\text{in}} - 0.541 \times 10^{-2}P_S + 0.090 \times 10^{-5}P_P - 30.302, \quad (16)$$

$$\text{WRG} = 1.93 \times 10^{-3}\rho - 0.011\nu + 0.292T_{\text{out}} - 281.11. \quad (17)$$

According to the outcomes of the following hypothesis tests, the MLR model for predicting the WRG is validated.

1. Analysis-of-variance test: $F_0 = 24.105$ and the corresponding $P_{\text{value}} = 0.000$.
2. Individual coefficient test: $P_{\text{value}} < 0.05$ (defined level of significance) for t_0 statistics of ρ, ν and T_{out}.
3. Non-multi colinearity test; variance inflation factors of ρ, ν and $T_{\text{out}} < 4$.
4. Durbin–Watson test: $D_0 = 2.247 > D_{\alpha,u} = 1.74$.

Based on the outcomes of these tests of hypothesis, the validity of the MLR model in predicting T_{out} was accepted. The results of the performance evaluation of the MLR model in predicting the granule PSI, based on the set of the test data, are given in Table 2.

5. Discussion and scenarios

The indicators used for evaluation of the model performance are MRE, RMSE and R^2. The comparison of numerical results for performance evaluation shows that the ANN model has an improvement in R^2 of 14.23%, a decrease in MRE of 6.72% and a decrease in RMSE of 2.27×10^{-3} kg when compared to the MLR model. These results indicate that the complexity and nonlinearity of the spray drying treatment may hardly allow a parametric model to discover and capture the relationships between the input and output variables.

In addition, according to the results of Table 2, the PC-ANN model has an improvement in R^2 of 5.84%, a decrease in MRE of 3.36% and a decrease in RMSE of 0.77×10^{-3} kg when compared to the ANN model. These results demonstrate

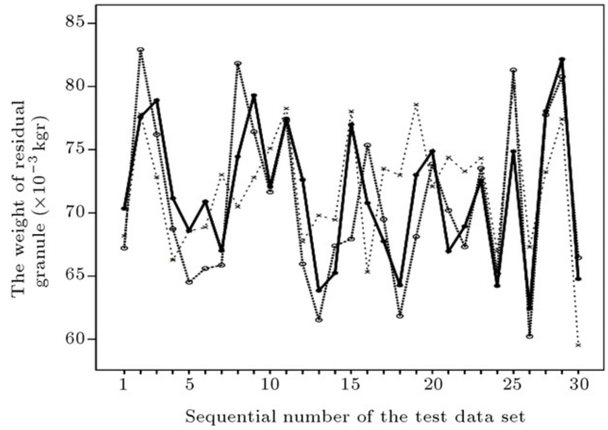

Figure 7: Comparison of predicted values of PC-ANN (♦) and ANN models (×) versus actual values (•) for the set of the test data.

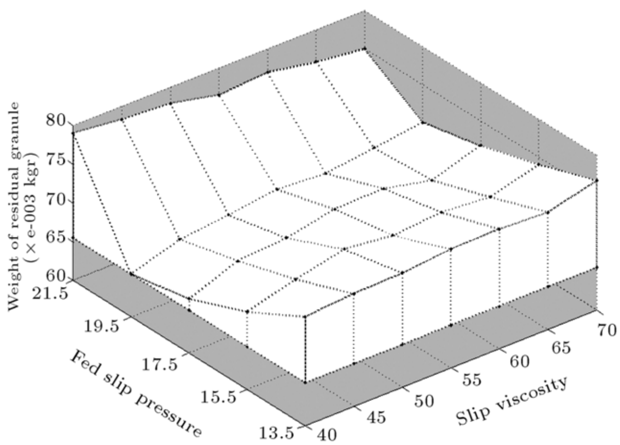

Figure 8: An illustration of several scenarios using predictions of the PC-ANN model for the process response as a function of the slip viscosity and the fed slip pressure.

that a proper dimension reduction in the input set, based on the findings of the PC analysis, leads to an improvement in the performance indicators of the PC-ANN model. In other words, the generalization capability of the PC-ANN model, via eliminating redundant inputs, tends to increase.

In order to compare the predicted values rendered by ANN and PC-ANN models versus the actual values of test data on an individual basis, one can refer to Figure 7. This figure shows that the PC-ANN predictions have a distinct visual correlation with actual values, even when the samples are selected beyond the normal range of actual values. For instance, even though the 3rd, 13th, 18th, 26th and 27th observations have been gathered out of the normal range of operating conditions, the PC-ANN model is capable of extrapolating them properly.

An important feature in modeling the spray drying mechanism is that the superior model, i.e. the PC-ANN model, can also be used to accurately investigate the effects of the input variables on the output. In order to visualize this capability, the PC-ANN model predictions of granule PSI, as a function of input variable, ν, for instance, are shown in Figures 8 and 9. These figures illustrate several scenarios for predictive process control when two input variables vary and the other input variables are kept constant. In fact, first, the values of T_{out} are calculated in terms of T_{in}, P_S and P_p, and then, they are used for prediction of WRG values via the PC-ANN model. Note that Figures 8 and

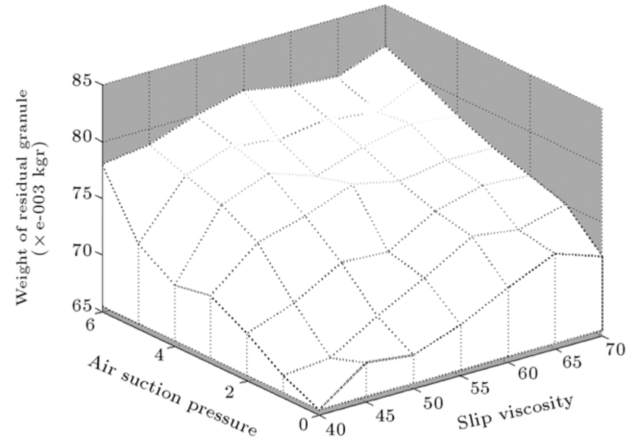

Figure 9: An illustration of several scenarios using predictions of the PC-ANN model for the process response as a function of the slip viscosity and the air suction pressure.

9 report the predictions not only in the normal range of inputs considered in the study, but also beyond the normal range.

Figure 8 depicts the changes in the predicted values of WRG, with respect to ν and P_p, when the other three input variables are kept constant at the values shown ($\rho = 1.68$, $P_S = 18.54$ and $T_{\text{in}} = 549.12$). As can be seen, in spite of what one expects to observe, WRG decreases, while P_p is more than 21. This abnormal behaviour can be interpreted as follows. When P_p increases significantly, a cloud of slip particles forms due to excessive height of spraying, which in turn, causes the slip particles to stick together. However, there is little exclusion, considering the interactive effects among variables.

Figure 9 reports the changes in the predicted values of WRG, with respect to ν and P_S, when the other three input variables are kept constant at the values shown ($\rho = 1.68$, $P_p = 17.30$ and $T_{\text{in}} = 549.12$). It can be seen that WRG increases, while P_S decreases with any increase in ν, because of a greater chance of particles sticking together. However, several exclusions considering the interactive effects among variables can be observed.

6. Conclusions

This study has been undertaken to present an integrated approach for prediction of granule particle size using Partial Correlation (PC) analysis and Artificial Neural Networks (ANN). According to the proposed algorithm, initially, a PC analysis is used to process actual data and to provide the necessary background for applying ANN. Next, with respect to the findings of the PC analysis, the PC-ANN model is developed for predicting the granule PSI. The performance of both models, based on their predictive capability, was evaluated, using such indicators as the coefficient of determination (R^2), Mean Relative Error (MRE) and Root Mean Square Error (RMSE). Comparing the performance of both models reveals that the PC-ANN model outperforms the ANN model in terms of an improvement in R^2 of 5.84%, a decrease in MRE of 3.36% and a decrease in RMSE of 0.77×10^{-3} kg. In fact, a proper reduction in the dimension of the input data set, based on a PC analysis, leads to improving the performance of the PC-ANN model. Besides, a comparison of the predictions of both models versus the actual values for a set of test data demonstrates that the PC-ANN model, not only interpolates more accurately than the ANN model, but also can extrapolate more satisfactorily.

Table A.1: A sample of raw data of spray drying in a large ceramic tile manufactory in Yazd province in south Iran.

Data number	1	2	3	4	5	6	7	···	292	293	294	295	296	297	298	299	300
$\rho(\times 10^3)$	1.66	1.68	1.69	1.69	1.66	1.67	1.67	···	1.67	1.68	1.68	1.67	1.66	1.68	1.68	1.67	1.69
ν	60	65	59	58	70	65	58	···	60	61	58	59	66	58	64	64	66
T_{in}	555	543	549	547	551	552	550	···	547	550	549	551	549	547	540	552	548
$P_p(\times 10^{-5})$	18.7	21.2	20.1	17.9	19.6	21.3	20.3	···	18.8	20.3	21.3	18.7	17.9	20.9	20.8	19.7	19.6
$P_S(\times 10^{-2})$	19	18.5	19	19.5	20	19	19.5	···	20.5	20.05	19.5	21	21	20.5	20	19	19
T_{out}	98	96	98	91	92	95	101	···	97	97	102	98	96	93	93	98	97
WRG $(\times 10^{-3})$	72	69.2	69.5	69.9	72.1	63.1	63.4	···	67.9	68.6	69.4	69.5	72.8	74.6	76.6	64.3	74.5

Finally, deploying the superior model, several scenarios, as accurate, fast running and inexpensive tools, are presented to identify the optimal process settings based on the desired process response. Production engineers employing the PC-ANN model, as a reliable model for predictive control of complex processes, can save both engineering effort and funds.

Appendix

The sample of spray drying raw data which was utilized in this experiment is presented in Table A.1. This set of actual data was gathered from one spray drier in a large ceramic tile manufactory in Yazd province, Iran.

References

[1] Anonymous, *From Technology Through Machinery to Kilns for SACMI Tile*, 2th ed., pp. 66–98, SACMI Press, Italy (1989).

[2] Youssefi, Sh., Emam-Djomeh, Z. and Mousavi, S.M. "Comparison of artificial neural network (ANN) and response surface methodology (RSM) in the prediction of quality parameters of spray-dried pomegranate juice", *Dry. Technol.*, 27(7), pp. 910–917 (2009).

[3] Barletta, M., Gisario, A. and Guarino, S. "Modeling of electrostatic fluidized bed (EFB) coating process using artificial neural networks", *J. Eng. Appl. Artifi. Intel.*, 20, pp. 721–733 (2007).

[4] Mandal, S., Sivaprasad, P.V., Venugopal, S., Murthy, K.P.N. and Raj, B. "Artificial neural network modeling of composition-process-property correlations in austenitic stainless steels", *J. Mater. Sci. Eng.*, 485, pp. 571–580 (2008).

[5] Jiang, Z., Gyurova, L., Zhang, Z., Friedrich, K. and Schlarb, A.K. "Neural network based prediction on mechanical and wear properties of short fibers reinforced polyamide composites", *J. Mater. Des.*, 29, pp. 628–637 (2008).

[6] Su, C. and Hsieh, K.L. "Applying neural network approach to achieve robust design for dynamic quality characteristics", *Int. J. Qual. Reliab. Manag.*, 15(5), pp. 509–519 (1998).

[7] Yimnirun, R., Laosiritaworn, Y.W., Khamman, O. and Ananta, S. "Artificial neural network modeling of ceramics powder preparation: application to $NiNb_2O_6$", *Ceram. Int.*, 34, pp. 809–812 (2008).

[8] Bonifazi, G. and Burrascano, P. "Ceramic powder characterization by multilayer perceptron (MLP) DATA compression and classification", *Adv. Powder. Technol.*, 5(3), pp. 225–239 (1994).

[9] Dong, G., Yongli, W., Juntao, X., Cewen, N. and Longtu, L. "Investigation of $BaTiO_3$ formulation: an artificial neural network (ANN) method", *J. Eur. Ceram. Soc.*, 22, pp. 1867–1872 (2002).

[10] Sukthomya, W. and Tannock, J. "The training of neural networks to model manufacturing process", *J. Intel. Manuf.*, 16, pp. 39–51 (2005).

[11] Afzulpurkar, N.V., Saeed, I.A. and Vyawahare, V.S. "Ceramic tile process modeling for quality improvement using ANN", *IEEE Int. Conf. on Indust. Technol.*, 15–17, Mumbai, India, pp. 2896–2900 (2006).

[12] Sterjovski, Z., Pitrun, M., Nolan, D., Dunne, D. and Norrish, J. "Artificial neural networks for predicting diffusible hydrogen content and cracking susceptibility in retile flux-cored arc welds", *J. Mater. Process. Technol.*, 184, pp. 420–427 (2007).

[13] Basheera, A.C., Dabadea, U.A., Joshia, S.S., Bhanuprasadc, V.V. and Gadre, V.M. "Modeling of surface roughness in precision machining of metal matrix composites using ANN", *J. Mater. Process. Technol.*, 197, pp. 439–444 (2008).

[14] Lam, S.Y., Petri, K.L. and Smith, A.E. "Prediction and optimization of a ceramic casting process using a hierarchical hybrid system of neural networks and fuzzy logic", *IIE. Trans.*, 32, pp. 83–91 (2000).

[15] Scott, D.J., Coveney, P.V., Kilner, J.A., Rossiny, J.C.H. and Alford, N.M.N. "Prediction of the functional properties of ceramic materials from composition using artificial neural networks", *J. Eur. Ceram. Soc.*, 27, pp. 4425–4435 (2007).

[16] Dinh, N.Q. and Afzulpurkar, N.V. "Neuro-fuzzy MIMO nonlinear control for ceramic roller kiln", *Simul. Model. Pract. Theory.*, 15, pp. 1239–1258 (2007).

[17] Yu, J., Wang, H. and Zhang, J.' "Neural network modeling and analysis of gel casting preparation of porous Si_3N_4 ceramics", *Ceram. Int.*, 35, pp. 2943–2950 (2009).

[18] Thyagarajan, T., Panda, R.C., Shanmugan, J., Rao, V.P.G. and Ponnavaikko, M. "Development of ANN model for non-linear drying process", *Dry. Technol.*, 15(10), pp. 2527–2540 (1997).

[19] Thyagarajan, T., Shanmugam, J., Panda, R.C., Ponnavaikko, M. and Rao, P.G. "Artificial neural networks: Principle and application to model based control of drying systems-A review", *Dry. Technol.*, 16(6), pp. 931–966 (1998).

[20] Kaminski, W., Strumillo, P. and Tomczak, E. "Neurocomputing techniques to modeling of drying process dynamics", *Dry. Technol.*, 16(6), pp. 967–991 (1998).

[21] Chegini, G.R., Khazaei, J., Ghobadian, B. and Goudarzi, A.M. "Prediction of process and product parameters in an orange juice spray dryer using artificial neural networks", *J. Food Eng.*, 84, pp. 534–543 (2008).

[22] Cui, B. and Han, R. "Modeling of dimensional errors in slender bar turning using artificial neural networks", *7th Int. Conf. on Intelligent Control and Automation*, 25–27, Chongqing, China, pp. 6053–6056 (2008).

[23] Demuth, H. and Beale, M., *Neural Network Toolbox User's Guide*, 4th ed., The Mathworks, Inc, MA, USA (2000).

[24] Krycha, K.A. and Wagner, U. "Applications of artificial neural networks in management science: a survey", *J. Retail. Consum. Serv.*, 6, pp. 185–203 (1999).

[25] Nascimento, C.A.O., Giudici, R. and Guardani, R. "Neural net based approach for optimization of industrial chemical process", *Comput. Chem. Eng.*, 24, pp. 2303–2374 (2000).

[26] Wilcox, J.A.D. and Wright, D.T. "Towards pultrusion process optimisation using artificial neural networks", *J. Mater. Process. Technol.*, 83, pp. 131–141 (1998).

[27] Sukhhomya, W. and Tannock, J. "The training of neural networks to model manufacturing process", *J. Intell. Manuf.*, 16, pp. 39–51 (2005).

[28] Islam, M.R., Sablani, S.S. and Mujumdar, A.S. "An artificial neural network model for prediction of drying rates", *Dry. Technol.*, 21(9), pp. 1867–1884 (2003).

[29] Morrison, D.F., *Multivariate Statistical Methods*, McGraw-Hill, New York, UK (1976).

Najmeh Neshat received her B.S. and M.S. Degrees in Industrial Engineering from the Departments of Industrial Engineering at Yazd University, Iran, in 2001 and Sharif University of Technology, in Tehran, in 2008, respectively. Currently she is a Ph.D. student of Industrial Engineering in the Department of Industrial Engineering at Tarbiat Modares University in Tehran, Iran. She has worked in the ceramic industry as consultant for over 6 years. Her current research interests are in Artificial Intelligence.

Aliyeh Kazemi was born in Iran, in 1980. She received her B.S. Degree in Industrial Management from the Department of Industrial Management, Faculty of Administrative Science and Economics, Isfahan University, in 2002, and her M.S. degree in Operations Research Management from the Department of Industrial Management, Faculty of Management, Tehran University, in 2005. Currently she is a Ph.D. student of Operations Research Management at the Department of Industrial Management, Faculty of Management, Tehran University.

Her employment experience includes the National Iranian Oil Refining and Distribution Company, Tehran, Iran (2004–2006) and the Petroleum Ministry, Tehran, Iran (2006–2009). Her current research interests include Operations Research, Energy Models and Artificial Intelligence.